Knowledge engineering for decision support in osteoporosis

Grant closeout final report

Principal Investigator:

Joanne LaFleur, PharmD, MSPH
Assistant Professor
University of Utah College of Pharmacy

Team Members:

Robert Adler, MD
McGuire VA Medical Center

Daniel Malone, RPh, PhD
Professor
University of Arizona College of Pharmacy

Jonathan Nebeker, MD, MS
George E Wahlen VA Medical Center
Professor of Internal Medicine
University of Utah School of Medicine

Organization:

University of Utah College of Pharmacy

Inclusive Dates of the Project:

2009 – 2015

Federal Project Officer:

Tamara Willis
Health Science Administrator
AHRQ/OERE/DRE

Acknowledgement of Agency Support:

This work was funded by the Agency for Healthcare Research and Quality

Grant Number:

K08-HS018582

STRUCTURED ABSTRACT

Purpose
Our goal was to produce a Veterans Affairs (VA)-optimized fracture absolute risk assessment (VA-FARA) rule for identifying males at highest risk of osteoporotic fracture, clinical decision rule, and decision support tool.

Scope
The goal was to use knowledge engineering to design decision support that incorporates clinicians' needs and minimizes cognitive burden.

Methods
We developed the rule using epidemiologic methods in national VA datasets of risk factors collected passively as a routine part of healthcare operations, then validated and calibrated it. We piloted it in the Veterans Integrated Service Network 21. We compared clinical strategies for the rule at different absolute risk thresholds, under a range of treatment efficacy assumptions. We evaluated those assumptions with a meta-analysis of randomized controlled trials. We identified implementation barriers by conducting focus groups among US clinicians.

Results
The rule performed with acceptable discrimination (C-statistic 0.7-0.8). In the VISN 21 pilot, our rule was superior for identifying highest-risk patients. In our comparison of clinical strategies, treating high-risk patients regardless of bone mineral density was most effective and least costly. Thresholds of absolute risk favoring specificity over sensitivity or accuracy were optimal. Our meta-analysis yielded no evidence of reduced bisphosphonate efficacy in non-osteoporotic men.

Key Words
Fracture, veterans, osteoporosis, absolute risk assessment, cost-effectiveness

PURPOSE

Our goal was to produce (1) a Veterans Affairs (VA)-optimized risk stratification rule for fracture, (2) a VA-optimized clinical decision rule, and (3) a decision support tool designed to facilitate appropriate prophylaxis of osteoporotic fractures. Despite the availability of effective treatments for men,[1] diagnosis, screening, and treatment rates remain abysmal. Following a low-trauma (fragility) hip fracture, only 6.9% of men received a diagnosis for osteoporosis within 1 year, only 0.0-2.8% received a bone mineral density (BMD) scan, and only 0.00-3.4% received treatment.[2-4]

Some have theorized that low diagnosis, screening, and treatment rates are driven by differences in guidelines, such as different methods for guideline development or variability in recommendations.[3] However, recent guidelines have achieved some consensus, at least for the United States (US). Both the National Osteoporosis Foundation (NOF) and Endocrine Society recommend universal bone densitometry in men over 70 along with treatment for osteoporotic men; osteopenic men should also be treated, as guided by fracture absolute risk assessment (FARA).[5]

Despite the clarity in newer guidelines, institutional adoption of the guidelines remains low. In the US, no organizations have adopted a policy of system-wide BMD screening in eligible men or use of FARA-guided screening or treatment. This reluctance may be due, in part, to the perceived budget impact of treating and monitoring the estimated 1 in 5 men and 1 in 3 women over 50 who would be eligible.[6] In the VA, the largest single healthcare provider for elderly men in the US, a case-finding approach is recommended that excludes universal densitometry or FARA.[7] Further, the National Committee for Quality Assurance (NCQA) has a low bar for monitoring postmenopausal osteoporosis and has no quality standards for assessing male osteoporosis. The Healthcare Effectiveness Data and Information Set (HEDIS) measure related to postmenopausal osteoporosis is defined as a woman with a prior fragility fracture receiving either treatment or a BMD scan, basically a "wait for fracture" approach without primary prevention.[8] There is no HEDIS measure for male osteoporosis, implicitly a "do nothing" approach.

FARA has become a key component for guiding treatment decisions in osteoporosis fracture prevention, in large part, because of the inability of BMD screening alone to identify 55% of women and 80% of men who have fragility fractures.[9] The NOF's treatment guideline recommends treatment in women and men with an absolute 10-year fracture risk of 3% for hip or 20% for any major fracture.[5] However, the risk stratification rule recommended by NOF performs poorly in men,[10] and one from Women's Health Initiative (WHI) investigators is designed only for women.[11] The Garvan nomogram,[12] a tool designed for either sex, has been found to have acceptable discrimination in men, but is not incorporated into the guideline.[10]

In addition to having limited application in men, perhaps the biggest limitation of the risk assessment approach in the clinical care setting is that it requires considerable clinician time and attention to observing and documenting clinical risk factors. A potential solution to this problem lies in the increasing use of electronic medical records (EMRs). EMRs can enable health systems to cue clinicians to consider osteoporosis risk in appropriate patients. EMR-based reminders compared to usual care have been shown to increase appropriate interventions in patients with prior fractures from 6% to more than 50%.[13] EMRs also represent a mechanism for passively collecting risk factor information, without increasing the clinician's cognitive burden. The objective of these studies was to determine whether structured data from electronic clinical and administrative datasets could be used to predict male veterans at highest risk for fracture with acceptable discrimination. If so, passively collected risk factor information could be used to create "smart" clinical reminders that identify high-risk patients, while minimizing the time requirement for providers.

SCOPE

This work comprises a series of experiments designed to implement an effective, inexpensive means of accurately diagnosing males at high risk for fracture in the clinical setting. First, we created and validated the VA-optimized algorithmic risk stratification rule (a predictive model) using computerized data to identify males at high risk for fracture. Next, to create a VA-optimized clinical decision rule about how and when to use the VA-FARA and how to treat patients, we used cost-effectiveness analysis to evaluate different FARA-based clinical decision-making strategies. We also conducted a systematic review and meta-analysis of clinical trials of bisphosphonates in men to better understand the evidence for bisphosphonate use in men and how those benefits might vary depending on clinical risk factors; this meta-analysis was designed to inform treatment decisions in the absence of fracture. Finally, we designed a decision support tool to facilitate appropriate prophylaxis of osteoporotic fractures and implemented this tool in the Veterans Integrated Service Network (VISN) 21. We then compared the performance of the VA-FARA to an electronic version of the World Health Organization's (WHO) FRAX (eFRAX) using BMI rather than BMD. In ongoing work, we are finishing analyses of

clinician focus group data that will help us refine the tool to provide information preferred by providers and conducting usability testing on the tool.

VALIDATED RISK RULE USING COMPUTERIZED DATA TO IDENTIFY MALES AT HIGH RISK FOR FRACTURE

In our efforts to develop computerized clinical decision support tools for identifying and treating male osteoporosis in Veterans, the first step was to create an algorithm that accurately predicts which males are at highest risk.

Methods

This study employed a cohort design using data from the VHA in the Veterans Integrated Service Network (VISN) 19, which covers much of the Rocky Mountain region of the United States. We accessed the data through the Austin Automation Center and our local data warehouse and constructed a dataset that included clinical, administrative, and utilization variables from the inpatient, outpatient, and pharmacy records.

We identified male veterans age 50 and older who received treatment in VISN 19 at any time during 2005 and 2006. An index date was defined as the first encounter after the latter of (1) January 1, 2005, (2) the patient's 50th birthday, or (3) 395 days after his first encounter with the VHA system. The latter requirement was to ensure a minimum duration of pre-index observation to identify baseline risk factor constructs. Because the goal of the program was to identify patients who would need treatment, patients were excluded if they had received an oral bisphosphonate on or before their index date. We excluded patients with missing or invalid body mass index (BMI) observations after iteratively determining that BMI was an important predictor of fracture in the data. Similarly, we decided to include patients with missing race information after confirming that race was not an important predictor of fracture. Patients were censored at the date they initiated treatment with an oral bisphosphonate, at the occurrence of a fracture endpoint for each of the models, or on December 31, 2007. Female patients were also excluded since we anticipated less than 5% of our veteran cohort would be female, which would not represent a large enough sample size to construct stable statistical models for women.

Table 1. Adjusted hazards and risk score "points" for predicting hip and any major fractures in final stratified Cox models (N=84,763)

Risk factor	N	Events	Hazards ratio[a]	95% CI	Risk score contribution (in points)
Hip fracture model					
Prior fracture	894	40	56.4	32.0-99.7	564
Age ≥80 years	11817	66	2.8	3.0-4.0	28
Normal weight vs. overweight	18761	92	2.3	1.7-3.1	23
Underweight vs. overweight	808	6	3.7	1.6-8.6	37
Complications of diabetes	8544	57	1.9	1.3-2.7	19
Malnutritive disorder[b]	636	9	2.6	1.3-5.1	26
Stroke	6514	37	1.5	1.0-2.2	15
Smoking	15943	58	1.3	1.0-2.2	13
Alcohol abuse disorder	8233	40	1.6	1.0-2.5	16
6-12 clinic visits in prior year (vs. ≤5)	23633	56	2.0	1.3-3.0	20
13+ clinic visits in prior year (vs. ≤5)	21727	128	3.7	2.5-5.5	37
Fall risk[c]	311	13	1.6	0.8-3.2	16
Any major fracture model					
Prior fracture	894	134	14.4	8.60-24.03	144
Age ≥80 years	11817	190	1.6	1.37-1.95	16
Normal or underweight vs. overweight	19569	300	1.3	1.14-1.54	13
Malnutritive disorder[b]	636	26	1.7	1.09-2.66	17
Opioid exposure	27217	564	1.3	1.14-1.52	13
Proton-pump inhibitor (PPI) use	22931	396	1.3	1.09-1.46	13
Depression diagnosis	21926	392	1.2	1.07-1.44	12
Stroke	6514	127	1.2	1.01-1.52	12
Seizure disorder	1539	152	1.6	1.00-2.58	16
Alcohol abuse disorder	8233	199	1.7	1.45-2.09	17
Fall risk[c]	311	145	1.7	1.33-2.05	17
6-12 clinic visits in prior year (vs. ≤5)	23633	277	1.3	1.08-1.50	13
13+ clinic visits in prior year (vs. ≤5)	21727	479	1.8	1.53-2.16	18

[a] *Adjusted for all other variables shown within each model*

[b] *Malnutritive disorder includes Kwashikor, nutritional marasmus, other and unspecified protein calorie malnutrition, intestinal malabsorption, and symptoms concerning nutrition, metabolism, and development*

[c] *Fall risk includes history of falls and gait instability*

We adapted potential clinical risk factor constructs from the FRAX and WHI fracture risk prediction algorithms. We also considered other factors that have known or theoretical associations with fracture in order to determine if those

constructs could improve identification of patients at highest risk. We queried clinical and administrative datasets to identify potential indicators of risk factors in the pre-index period and considered multiple constructions of covariates for predicting risk. For example, we considered both a dichotomous indicator of past glucocorticoid exposure as a risk factor and contrasted that with a categorical variable for different degrees of cumulative glucocorticoid exposure in the pre-index period.

Table 2. Test characteristics for the hip and any major fracture risk stratification rules at 3 optimal cut-offs in all patients, in the subset with a prior fracture, and in the subset with no prior fracture

Hip fracture									
Characteristic	**Absolute 10-year risk[b] 1.8% (34 points) Favors sensitivity**			**Absolute 10-year risk[b] 2.5% (39 points) Favors accuracy**			**Absolute 10-year risk[b] 4.0% (62 points) Favors specificity**		
	All	**Prior fracture**	**No prior fracture**	**All**	**Prior fracture**	**No prior fracture**	**All**	**Prior fracture**	**No prior fracture**
Mean accuracy[c]	0.70	n/a[d]	0.69	0.72	n/a[d]	0.69	0.70	n/a[d]	0.65
Proportion classified correctly	0.57	0.04	0.57	0.70	0.04	0.71	0.88	0.04	0.88
Sensitivity	0.84	1.00	0.80	0.73	1.00	0.67	0.52	1.00	0.42
Specificity	0.57	n/a[d]	0.57	0.70	n/a[d]	0.71	0.88	n/a[d]	0.89
PPV (absolute 33-month risk)	0.005	0.045	0.004	0.007	0.045	0.005	0.011	0.045	0.008
NPV	0.999	n/a[d]	0.999	0.999	n/a[d]	0.999	0.999	n/a[d]	0.999
Any major fracture									
Characteristic	**Absolute 10-year risk[b] 7.3% (31 points) Favors sensitivity**			**Absolute 10-year risk[b] 9.5% (44 points) Favors accuracy**			**Absolute 10-year risk[b] 12.4% (56 points) Favors specificity**		
	All	**Prior fracture**	**No prior fracture**	**All**	**Prior fracture**	**No prior fracture**	**All**	**Prior fracture**	**No prior fracture**
Mean accuracy[c]	0.66	n/a[d]	0.64	0.67	n/a[d]	0.64	0.66	n/a[d]	0.62
Proportion classified correctly	0.57	0.150	0.57	0.73	0.15	0.57	0.84	0.15	0.84
Sensitivity	0.75	1.00	0.71	0.62	1.00	0.71	0.48	1.00	0.40
Specificity	0.57	n/a[d]	0.57	0.73	n/a[d]	0.57	0.84	n/a[d]	0.85
PPV (absolute 3-year risk)	0.020	0.150	0.017	0.026	0.150	0.017	0.034	0.150	0.026
NPV	0.995	n/a[d]	0.995	0.994	n/a[d]	0.995	0.993	n/a[d]	0.993

Key: PPV – Positive predictive value; NPV – Negative predictive value; n/a – Not applicable.

[a] Some possible cut-points were selected based on maximal and optimal mean accuracies.

[b] Absolute 5- and 10-year risks are estimated using the assumption that 1-year risk is equivalent to one-tenth of 10-year risk.

[c] Mean accuracy equals mean of sensitivity and specificity.

[d] Mean accuracy, specificity, and NPV could not be calculated for the subsets of patients with a prior fracture because all optimal cut-points classified all patients with a prior fracture as being at risk for future fractures.

Outcomes of interest were (1) the occurrence of a hip fracture or (2) the occurrence of a fracture of the hip, spine, forearm, and proximal humerus. These sites are the basis for the 10-year absolute fracture risk estimates included in the FRAX. [14] Diagnosis codes from the 9th revision of the International Classification of Diseases (ICD-9) were used to identify fracture outcomes including 820-821 for hip, 805 for vertebral, 813 for forearm, and 812 for proximal humerus fractures. Our focus here was to identify predictors of fragility fractures. Therefore we excluded fracture codes that occurred with a trauma code (E-code).

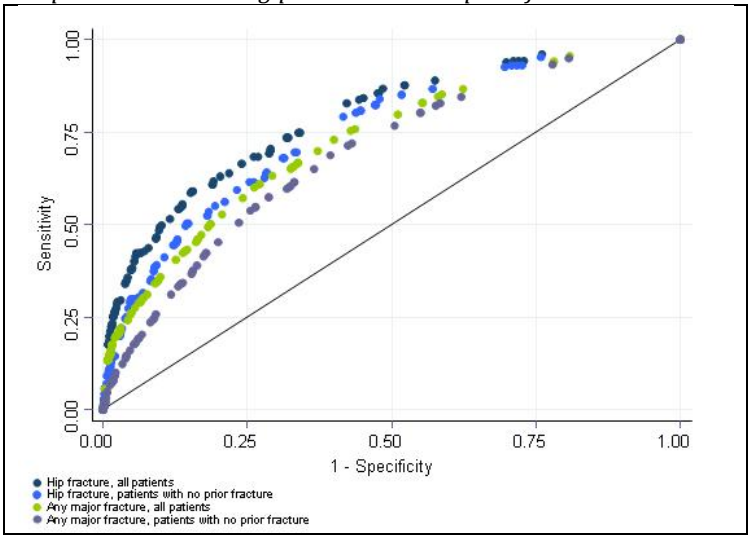
The risk stratification rule was generated in a multi-step process. First, Cox Proportional Hazards models were used to estimate the risk of hip and fragility fracture associated with each risk factor in univariate and multivariable analyses for all patients and for patients with no prior fracture separately. This is because patients with a prior fracture are eligible for an intervention according to all treatment guidelines in place at the time of the analysis, [5,15,16] and because we anticipated that the 'prior fracture' variable would overwhelm the models and cause important risk factors among patients with no prior fracture to fall out. Multivariable models were constructed using a forward stepwise procedure, selecting candidate risk factor constructs from those that were significant in univariate analyses at an *a priori* significance level of 0.1. Variables were retained in the models if they were statistically significant at an alpha level of 0.05 or modified the association between other risk factors and the fracture outcome by more than 10%. A backward stepwise procedure was also used to confirm the stability of predictors identified in the forward procedure. We then selected the prior fracture variable and all the variables that remained in the model for the 'no prior fracture' cohort and

put them all into a single multivariable model, stratified by prior fracture status, to get accurate adjusted hazards ratios for each could be estimated in the overall cohort.

We used the C-statistic (concordance index), or area under the Receiver Operator Characteristic (ROC) curve, to validate the final Cox regression models. Since prognostic models tend to be overly optimistic (either because of regression-to-the mean bias or overfitting), we estimated the bootstrap-resampled honest C-statistic (or shrunken C). [17,18] This involved the re-estimation of test characteristics on 200 samples the size of our cohort, drawn with replacement. The bootstrap resampled C-statistic represents what we would expect to observe after regression to the mean has occurred or when overfitting was not present. Although split-sample techniques have been commonly used historically, we chose not to use that approach since the training set would be statistically equivalent to the test set, and consequently that approach would underestimate the optimism of the model for future populations. Bootstrap validation tests the performance of the model on numerous permutations of the population that differ slightly in risk factor characteristics, and it is currently regarded as a better approach to validation than split-sample techniques. [18-20]

Risk contributions ("points") for each risk factor were created by multiplying adjusted hazards ratios from the final models by 10 and rounding them to the nearest integer. Patient-level risk scores were then constructed by adding the contributions from each risk factor that was present at baseline. Finally, we used logistic regression to predict fractures with the risk score as the independent variable for each outcome. Test characteristics for the rules were calculated at various potential cut-points including percentage classified correctly, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and C-statistic (mean accuracy or area under the ROC curve). The rules were calibrated by comparing the proportions of events that were observed versus expected across the range of risk levels. All analyses and calculations were conducted using Stata SE v. 10 (StataCorp, College Station, Texas) and SAS v. 9 (Cary, NC).

Figure 1. Receiver operator characteristic (ROC) curves for hip fracture and any major fracture risk stratification rules among all patients and among patients with no prior fracture



Results

Among the population of 115,012 patients older than age 50 who received care in 2005 or 2006, 3,072 (2.7%) patients had been treated with an oral bisphosphonate prior to the index date and 27,177 (23.6%) patients missing height and/or weight were excluded, leaving 84,763 in the analysis. The mean age of the cohort was 66.9 (SD 10.3). The mean BMI was 28.8 (SD 5.3). Race or ethnicity data was only present in 42.5% of the cohort; of those for whom race data was available, 87.8% were Caucasian, 6.5% were Hispanic, and 4.5% were black.

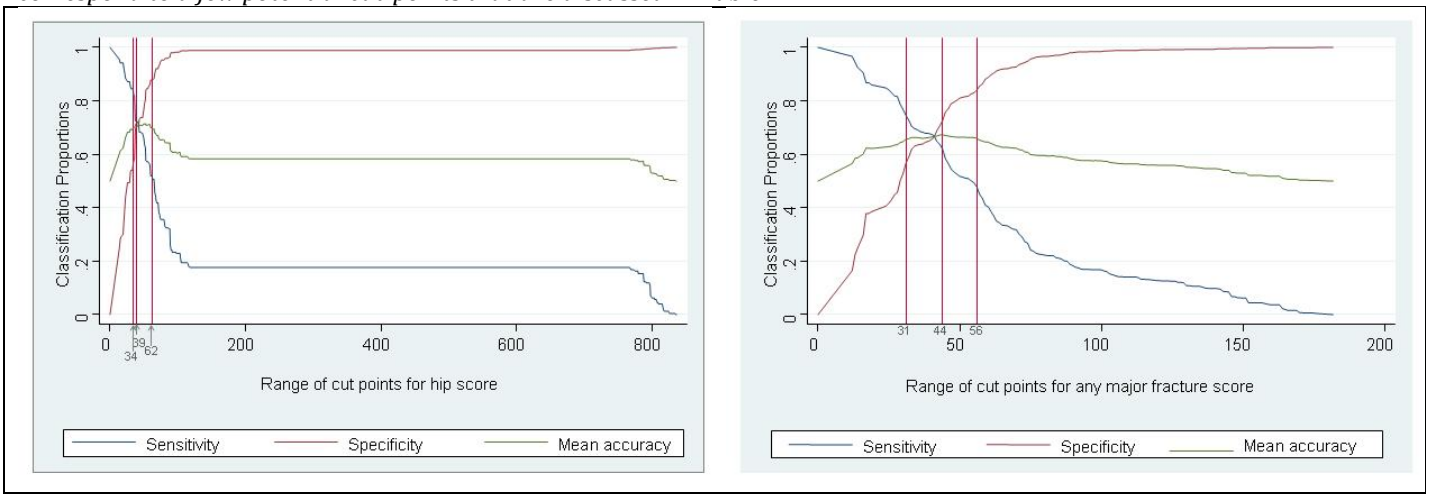
There were 227 non-traumatic, non-pathologic hip fractures over an average of 33.3 months. This corresponds to an incidence rate of 1.0 hip fractures per 1000 person-years. Among those with a hip fracture, the average time-to-event was 17.9 months. For non-traumatic, non-pathologic fractures of the hip, spine, forearm, or proximal humerus (any major fracture), 987 occurred over an average follow-up time of 33.2 months, corresponding to an incidence rate of 4.3 fractures per 1000 person-years. Among those with any major fracture, the average time-to-event was 21.3 months.

A summary of the risk estimates from the final multivariable Cox models along with the assigned contributions of each risk factor to the overall risk scores is given in Table 1. As we suspected, several risk factors that were highly predictive of fracture in the subset of patients without a prior fracture did not significantly predict fracture in the full cohort and fell out of predictive models generated using stepwise procedures. This supported our decision to use final Cox models that were stratified on prior fracture status. For example, in the hip fracture models, the risk factors for diabetes complications, malnutritive disorders, history of stroke, and smoking-related disorders were important and highly significant in the subset of patients without a prior fracture (data not shown). However, they were non-significant in the overall cohort, most likely due to the overwhelming effect of prior fracture in those models.

The C-statistics for the Cox models were 0.81 for hip fracture and 0.74 for any major fracture, suggesting the regression models had acceptable to good discrimination. In both models, the bootstrap-estimated (honest) C-statistic was similar (0.82 for hip fracture and 0.74 for any major fracture), suggesting that there was little or no optimism in the models. This finding was also consistent in the subset of patients without a prior fracture. In that subset, the original and honest C-statistics were the same both for hip fracture (0.78) and any major fracture (0.70).

For the hip fracture endpoint, the risk stratification rule showed that each 1-point increase in score was associated with a 0.42% increase in odds of having an event in the logistic model (95% CI 0.38-0.46%). A cut-point for this rule where the mean accuracy (average of sensitivity and specificity) was maximized was 39, as shown in Table 2. At this cut-point, the risk stratification rule predicted 165 of the 227 hip fractures (73%) and missed 62 (27%). Others could also be chosen to balance the trade-off between sensitivity and specificity. At two such cut-points, 34 and 62, the rule's sensitivity and specificity ranged from 0.84-0.52 (sensitivity) and 0.57-0.88 (specificity). Because the score for a prior fracture (564 points) was above all optimal thresholds, all patients with a prior fracture who had a subsequent fracture (N=40) were identified by the model. The remaining 125 of the 165 identified fractures (75%) occurred in patients without a prior fracture. This represents 67% of the fractures that occurred in patients without a prior fracture. Overall, the rule had acceptable discrimination, defined as a C-statistic between 0.7 and 0.8, in both the overall cohort (C=0.79) and among those without a prior fracture (0.75). ROC curves for the risk stratification rules are shown in Figure 1 and sensitivity, specificity, and mean accuracy curves are shown in Figures 2a and 2b.

Figure 2. Sensitivity, specificity, and mean accuracy (i.e., mean of sensitivity and specificity; equivalent to C-statistic or area under the ROC curve at each point) for (a) hip and (b) any major fracture across the range of cut-points for each risk stratification rule. The x-axes correspond to the range of scores (potential cut-points) for each rule. The y-axes correspond to value for the sensitivity (blue), specificity (red), and mean accuracy (green) of the rule at each cut-point. The red vertical lines correspond to a few potential cut-points that are discussed in Table 2.



For the major fracture rule, the risk stratification rule showed that each 1 point increase in score was associated with a 2.7% increase in the odds of having a fracture of the hip, spine, forearm, or proximal humerus (95% CI 2.5-2.9%). A cut-point that maximized mean accuracy for the logistic model of the risk stratification rule for major fracture was a score of 44. At this cut-point, the risk stratification rule predicted 611 of the 987 major fractures (62%) and missed 376 (38%). Other cut-points were 31 and 56, where sensitivity of the rule ranged from 0.75 to 0.48 and specificity from 0.57 to 0.84. Because the score for a prior fracture (144 points) was above all optimal thresholds, all 134 patients with a prior fracture who had a subsequent fracture were identified by the model. The remaining 477 of the 611 identified fractures (78%) occurred in patients without a prior fracture, representing 56% of the major fractures that occurred in patients without a prior fracture. Overall, the rule had acceptable discrimination in the overall cohort (C=0.73). Among patients with no prior fracture, the rule approached acceptable discrimination (C=0.69).

In our examination of the observed versus expected events across the range of values for the rules, the calibration curves converged across the entire range for the any major fracture rule and across most of the range for the hip fracture rule. At the low-risk end of the curve for the hip fracture rule there was minor separation, which was attributable to low numbers of events in low-risk patients.

Summary
We were successful in developing fracture absolute risk assessment (FARA) rule that identifies males at highest risk for fracture with acceptable discrimination. A next step was to incorporate the FARA rule into a computerized clinical decision support tool to aid clinicians in making treatment decisions.

COST-EFFECTIVENESS STRATEGIES FOR OSTEOPOROSIS SCREENING AND FRACTURE PREVENTION USING FRACTURE ABSOLUTE RISK ASSESSMENT

Our goal was to use cost-effectiveness analysis to evaluate different FARA-based clinical decision-making strategies.

Methods

We created and tested an osteoporosis cost-effectiveness model using the perspective of the VA. We adapted a model first developed by Ito and colleagues[21,22] and implemented it using decision analysis software (TreeAgePro 2013 Suite, release v2.0, TreeAge Software, Inc., Williamston, MA). The screening strategies compared are described in Table 3. The model was a microsimulation capable of representing both first- and second-order uncertainty. It represented 4 health states, as illustrated in Figure 3. All patients started well in the community and remained there for at least 1 cycle (1 year). During that year, and depending on the treatment strategy, they experienced or did not experience transition states (fracture events, treatment adverse effects, death). At the end of each year, they either remained in the same health state or transitioned to a health state with lower functionality in a unidirectional manner. Transition states and treatments modified a patient's probability of progressing through the health states. Patients also progressed through health states for reasons unrelated to osteoporosis. Each patient's underlying probability of transitioning between health states was modified by their prior experience (e.g., the presence of underlying osteoporosis, background absolute fracture risk, or the occurrence of fractures), recorded by the use of tracker variables in TreeAge.

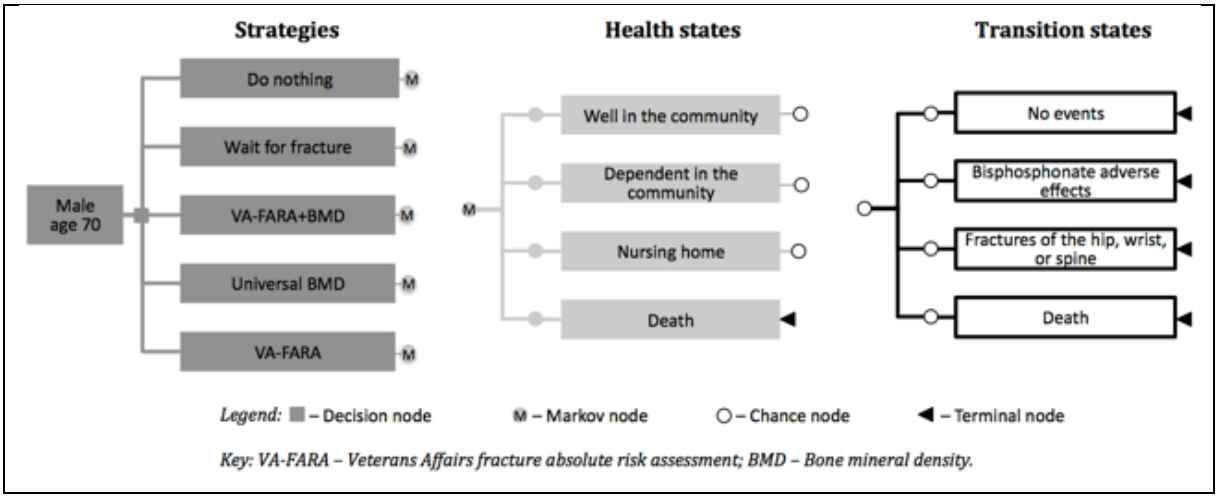
Table 3. Description of strategies compared in the cost-effectiveness analysis

Do-nothing	No antifracture interventions of any kind would be made, either screening or treatment.
Wait-for-fracture	No fracture interventions would be made unless a patient fractured, in which case the patient would receive bone densitometry and oral bisphosphonate treatment (a postmenopausal osteoporosis HEDIS policy).
Universal-BMD	All men over 70 would receive bone densitometry and those in the osteoporotic range would receive oral bisphosphonate (an older, pre-FARA policy).[23]
VA-FARA+BMD	FARA would be done (using VA-FARA) [24,25] and those considered high-risk would receive bone densitometry followed by bisphosphonates in osteoporotic patients (adapted from Britain's National Institute for Health and Care Excellence (NICE) guideline to reflect US preferences). [26]
VA-FARA	Assumed that absolute risk assessment would be done (using VA-FARA)[24,25] and those considered high risk would receive bisphosphonate treatment even without bone densitometry (suggested by the NICE guideline [26] and supported by the fact that 80% of men with fragility fractures have BMD T-scores outside the treatment range and would thus be untreated if treatment depended on bone densitometry in the osteoporotic range [9]).

Simulated patients

All patients entered the model at age 70; all were subjected to each treatment strategy simultaneously and transitioned through the model until death or 100 years of age. The model assumed that 8.1% of patients had underlying osteoporosis.[27]

Figure 3. Schematic representation of the Markov model.



Patients who were non-osteoporotic at baseline could subsequently become osteoporotic in 5-year increments up to a

prevalence of 14.3%.[27] Similarly, patients were also classified as high risk or not based on the presence of clinical risk factors and absolute fracture risk.[24] At model entry, the model assumed that 18.5% of patients had an absolute fracture risk that exceeded one of these thresholds, and this increased in 5-year increments up to a prevalence of 31.4%.[24]

Transition probabilities

Background transition probabilities for movement between health states are summarized in an electronic Appendix B (available from the PI upon request), including changes in background risk associated with treatment, risk factors, and transition states. Population fracture incidence rates were adjusted downward to account for the heterogeneous mix of risk factors in the population.[28] Mortality rates were derived from 2007 US life tables with excess mortality modeled following a hip fracture.[29] For efficacy, we used the pooled results from a meta-analysis of clinical trials in men, which reported a strong protective effect for vertebral fractures (odds ratio [OR] of 0.36) and a slightly less protective effect for non-vertebral fractures (OR 0.73).[30] We assumed that bisphosphonate efficacy would not take effect until 2 years after initiating treatment,[30] and the reduction in the relative risk of fracture remained as long as the patient continued on bisphosphonate therapy.[31] If a patient experienced an adverse event, they discontinued bisphosphonate therapy and the protective effect of therapy declined back to baseline linearly over 5 years. [31] Bisphosphonate therapy was not reinitiated in those patients.

Cost and utility inputs for the model are also summarized in an electronic Appendix B (available from the PI upon reques)t. All costs were standardized to 2013 US dollars using the consumer price index.[32] We included the annual cost of oral bisphosphonate (alendronate) in the VA, along with biannual BMD testing while on therapy. Costs incurred from fracture events included only the first year of fracture care. We compared costs (in 2013 US dollars), unadjusted patient life years, and quality-adjusted life years (QALYs) among the six strategies. We calculated the incremental cost-effectiveness ratio (ICER) for each strategy. Future costs and benefits were discounted at a 3% annual rate to reflect time preferences for costs and benefits.

Analysis and validation

The model was implemented using 50,000 first-order Monte Carlo microsimulations. Results were calculated in 5-year time periods up to age 90 and 10 years up to 100. Because fracture events are rare, the cycle length was set at 1 year with a half-cycle correction. From the model outputs, the overall incidence of fractures was calculated. Hip, vertebral, and wrist fracture incidence rates were compared to Melton's reported rates from a 1989-1991 Minnesota cohort, the most recent epidemiology study of all relevant clinical fracture types among males in the US,[33] as well as Ettinger's updates to Melton's numbers, based on 2006 hospital discharge data for hip fractures and based on modeled revisions to the Minnesota numbers for vertebral and wrist fractures. [34]

Sensitivity analyses

We conducted a deterministic sensitivity analysis in which half efficacy was assumed for bisphosphonates in non-osteoporotic men classified as high risk using FARA. This is because there is a lack of evidence that non-osteoporotic men who are otherwise high risk will receive as much benefit from bisphosphonate therapy as osteoporotic men, so FARA-guided treatment may be deemed as controversial despite being codified into treatment guidelines. [5,35] We also conducted a sensitivity analysis in which no efficacy was assumed for bisphosphonates in non-osteoporotic men classified as high-risk using FARA.

We carried out a probabilistic sensitivity analysis (PSA) to test uncertainty in parameter inputs using beta distributions for probabilities, log-normal distributions for relative risk variables, gamma distributions for costs, and beta distributions for utilities and utility decrements. The ranges for the PSA are summarized in the electronic Appendix B (available from the PI upon request). The PSA was carried out using 1,000 second-order parameter samples randomly selected from the distributions and 10,000 first-order trials per analysis. Results are displayed using a cost-effectiveness acceptability curve, representing the proportion of iterations where each screening method was considered cost-effective across a range of willingness-to-pay thresholds.

Results

Model validation

The simulation model produced overall fracture incidence rates for the do-nothing strategy of 4.86, 2.53, 3.44, and 0.77 per 1000 person-years for hip, clinical vertebral, sub-clinical vertebral, and wrist fractures, respectively, in men ages 70+. Ettinger did not report overall rates for the same age and fracture categories; however, he did report a hip fracture rate for a slightly different age cohort (ages 50+) of 2.09 per 1000 person-years, which compares plausibly with the higher

incidence rate we found for the age 70+ subset.[34] For age-specific incidence rates, our simulated rates were comparable to published population estimates.[33,34] The most extreme differences were observed for hip fracture in the 75-79 age category, for which the simulated rates were 23% higher (4.94 versus 4.02), and clinical vertebral fracture in the 85+ age category, for which the simulated rates were 13% lower (10.73 versus 12.39).[33,34]

Base-case analysis

Results from the base-case analysis are shown in Table 4 and Figures 4 and 5. VA-FARA dominated all other strategies, meaning that it was less costly and more effective. It as associated with an incremental cost savings per patient of $127, $230, $294, and $526 compared to do-nothing, wait-for-fracture, VA-FARA+BMD, and universal-BMD, respectively. The additional life years per patient for VA-FARA were 0.027, 0.023, 0.018, and 0.015, respectively, and the additional QALYs per patient were 0.017, 0.015, 0.015, and 0.010.

Table 4. Summary of outcomes in the base case: average cost, life years, and QALYs per patient

Strategy	Average per patient			Incremental difference vs. VA-FARA		
	Cost ($)	Life years	QALYs	Cost	Life years	QALYs
VA-FARA	169,989	17.796	10.944	-	-	-
Do-nothing	170,116	17.769	10.927	127.24	-0.027	-0.017
Wait-for-fracture	170,219	17.773	10.930	229.70	-0.023	-0.015
VA-FARA+BMD	170,283	17.778	10.932	293.96	-0.018	-0.012
Universal-BMD	170,516	17.781	10.934	526.76	-0.015	-0.010

Sensitivity analyses

With the half-efficacy assumption (where bisphosphonate efficacy was assumed to be reduced by 50% in non-osteoporotic men classified as high-risk using FARA), the VA-FARA strategy remained dominant over the wait-for-fracture, VA-FARA+BMD, and universal-BMD strategies. With the non-efficacy strategy, strategies incorporating VA-FARA were not preferred.

Figure 4. Average cost (2013 US dollars) and effectiveness (QALYs) per patient for different strategies in the base case

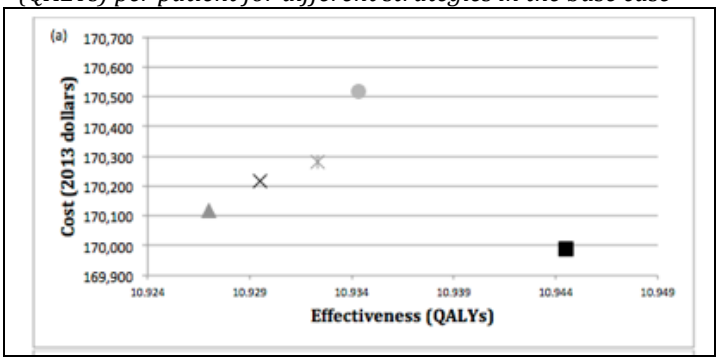
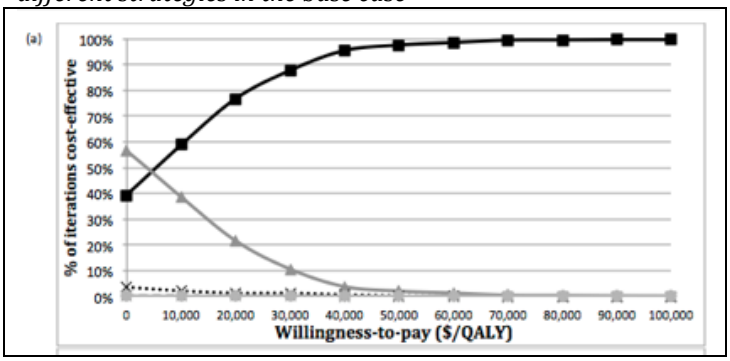


Figure 5. Cost-effectiveness acceptability curve for the different strategies in the base case



Summary

Our findings suggest that VA-FARA, a strategy characterized by treating on the basis of absolute risk alone, was less costly and more effective than all other strategies. However, we recognize that there is substantial clinical inertia for implementing that strategy. Causes of resistance to the strategy might include low clinical acceptance of treating non-osteoporotic, high-risk patients. An important next step for us was to address the uncertainty about bisphosphonate efficacy in these patients.

BISPHOSPHONATES FOR FRACTURE PREVENTION IN MALES: A SYSTEMATIC REVIEW AND META-ANALYSIS

Results from our cost-effectiveness analysis were dependent upon the efficacy of bisphosphonates in non-osteoporotic, high-risk patients, but that efficacy is an unknown entity. In addition, focus groups revealed that clinicians have considerable uncertainty about what to do about osteoporosis risk in patients, particularly in males, even when presented with considerable information about the patient's clinical condition (see under OTHER/ONGOING ANALYSES, below). One aspect of this problem is the lack of published data on the efficacy of bisphosphonates in male patients, and whether male patients with different clinical risk factors, who are otherwise high risk, would benefit from treatment. To better understand the evidence for bisphosphonate use in men, and to understand how those benefits might vary depending on clinical risk factors, we conducted a systematic review and meta-analysis of clinical trials of bisphosphonates in men. Our primary goals were to (1) quantify the efficacy of bisphosphonates in males at risk for osteoporosis fracture, and (2) to quantify the variability in efficacy associated with baseline risk factor information such as male sex, BMI, BMD, and age.

Methods

We selected studies based on 5 criteria: (1) we only accepted randomized controlled clinical trials (RCTs) with a duration of at least 1 year; (2) we only accepted RCTs with male participants at risk for fracture, including RCTs that contained both men and women if the number of males included in the study was reported separately; (3) we only accepted RCTs that used bisphosphonates at doses approved by the Food and Drug Administration (FDA) for treatment or prevention of osteoporosis; (4) the number of fractures at the end of the study must have been reported. Fracture types considered were any fracture, vertebral fractures, non-vertebral fractures, and hip fractures.

We searched PubMed/MEDLINE, Scopus, the Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library, Clinicaltrials.gov, and Embase. The search strategy generated by the Cochrane Musculoskeletal Group was used together with other specific key words in this study for MEDLINE and was adapted for the rest of the above electronic databases. All searches were conducted without language restrictions on October 20, 2014. The PubMed/MEDLINE search strategy is summarized in an electronic Appendix C (available from the PI upon request).

Study screening and selection was conducted in two phases: title/abstract review followed by full-text review. Title/abstract review was conducted after removing duplicates. Two reviewers independently examined titles and abstracts to identify eligible studies. A study was included for further review if two reviewers independently could not rule out that the study met the eligibility criteria listed above. Studies that seemed to meet the criteria were pulled for full-text review. Again, two reviewers independently reviewed each study to make sure it met the above.

Data extraction

Two reviewers independently extracted all required data from each study using a standardized form, which covered: (1) study characteristics (publication details, location, duration, intervention, and dosage); (2) baseline patient characteristics (number of patients, age, male proportion, average BMI, average %-score at lumbar spine, femoral neck and total hip, and prior fractures); and (3) outcomes (type of fracture, fracture definition, screened population, male population, number of patients with fracture, number of males with fracture. We considered all reported fractures.

The reviewers followed the instructions suggested by the Cochrane Handbook for Systematic Reviews of Interventions.[36] When there was disagreement in data extraction, both reviewers referred back to the original article and established a consensus. Several studies reported fracture outcomes in ways that could not be extracted without additional information. The original authors of these studies were contacted for further details, but the studies were excluded from the analysis if these authors did not respond or could not provide further data.

Analysis

Each included study's characteristics, including location, participants, duration, interventions, male proportion, mean BMIs and T-scores, prior fractures, and fracture outcomes were reported descriptively and qualitatively. The methodologic quality of the included studies was assessed according to the criteria specified by the Cochrane Collaboration's tool for assessing risk of bias.[36]

We calculated the relative risk (RR) of all reported fracture outcomes (i.e., any, clinical vertebral, morphometric vertebral, non-vertebral, hip) with 95% confidence intervals (CI) for each study. Data were pooled using a fixed effects model. The pooled RRs were calculated using the general inverse variance method for the weights, as described in Fleiss.[37] Statistical heterogeneity was assessed using a chi-square test with significance set at a *p*-value of 0.10 and the I^2

Table 5. Primary analysis results, male-only corpus

Outcome	# studies	# participants	Relative risk	95% CI
End of study				
Any fracture	19	2397	0.56	(0.36, 0.85)
Morphometric vertebral fracture	15	2143	0.40	(0.22, 0.73)
Clinical vertebral fracture	8	1872	0.75	(0.28, 2.04)
Non-vertebral fracture	13	1957	0.75	(0.37, 1.49)
Hip fracture	2	82	0.30	(0.01, 7.06)
12 months				
Any fracture	13	806	0.89	(0.46, 1.70)
Morphometric vertebral fracture	9	1579	0.48	(0.24, 0.95)
Clinical vertebral fracture	5	346	0.98	(0.27, 3.53)
Non-vertebral fracture	8	366	1.39	(0.16, 11.81)
Hip fracture	NA	NA	NA	NA
24 months				
Any fracture	8	1875	0.45	(0.29, 0.69)
Morphometric vertebral fracture	7	1834	0.30	(0.16, 0.57)
Clinical vertebral fracture	2	1483	0.56	(0.09, 3.55)
Non-vertebral fracture	4	1548	0.67	(0.31, 1.47)
Hip fracture	1	55	0.30	(0.01, 7.06)

Key: CI – confidence interval; NA – not available

test with substantial heterogeneity defined as I^2>50%.[38] Publication bias was examined using funnel plots and Egger's test.[39] The RCTs included in our analysis varied in terms of study duration and fracture outcomes reported. Additionally, some studies included males but did not report fractures separately for males. For these reasons, we performed analyses on the subset of studies that reported male outcomes separately (male-only corpus), as well as on all studies (male/female corpus). Our primary analyses examined separately the effect of bisphosphonates on each fracture outcome at the end of the study period and also at 12 and 24 months.

We conducted meta-regressions to examine the impact of the studies' male proportions on the effect size of bisphosphonates on prevention of any fracture and morphometric vertebral fracture in the male/female corpus. We conducted further meta-regression to examine the impact of age; BMI; and BMD T-score at lumbar spine, femoral neck, and total hip sites on the effect size of bisphosphonates on prevention of any fracture and morphometric vertebral fracture in the male-only corpus. A two-sided p-value ≤0.05 was considered statistically significant. All analyses were preformed using Stata software (Stata, 13.1; StataCorp LP, College Station, Texas, USA).

Results

Literature search

The searches of PubMed/MEDLINE, Scopus, CENTRAL, Clinicaltrials.gov, and Embase yielded 1325 abstracts, of which 363 were duplicates. A total of 962 abstracts were included in the Phase I title/abstract review. At Phase I, 807 articles were excluded, mostly because they were not an RCT (403) or did not study one of the FDA-approved bisphosphonates (177). There were 155 articles included in the Phase II full-text review. We excluded 112 of these studies because they did not report fracture outcomes (50), used non-licensed dosages (16), were not RCTs (9), used vitamin D or calcium in one group but not the other (5), had the same population as another included study (13), had less than 12 months of study duration (3), had un-extractable results (5), or were duplicates (11). As a result, we extracted data from 43 studies for inclusion in the male/female corpus, 19 of which were included in the male-only corpus.

Study characteristics

The characteristics of the 43 included studies are summarized in an electronic Appendix D (available from the PI upon request). The most common intervention, alendronate, was used in 24 studies, while 12 studies used risedronate, 5 used zoledronic acid, and 2 used ibandronate. Morphometric vertebral fractures were reported in 27 studies, clinical vertebral fractures were reported in 11 studies, non-vertebral fractures in 27 studies, and hip fractures in 10 studies. Most studies had a duration of 12 or 24 months.

Risk of publication bias

We found no evidence of publication bias based on funnel plots. The Egger test showed a *p*-value of 0.119 for the male/female corpus and of 0.690 for the male-only corpus, confirming no statistical evidence of publication bias.

Effects of bisphosphonates on fracture

The results of the pooled analysis for the male-only corpus are summarized in Table 5. When we pooled the data, using "end of study" as the time-point, bisphosphonates were shown to significantly reduce the risk of any fracture (RR = 0.56, 95% CI 0.36, 0.85). Bisphosphonates were also shown to significantly reduce the risk of morphometric vertebral fractures (RR = 0.40, 95% CI 0.22, 0.73) in the male-only corpus. Bisphosphonates were not shown to have a significant effect on clinical vertebral fracture (RR = 0.75, 95% CI 0.28, 2.04) or non-vertebral fracture (RR = 0.75, 95% CI 0.37-1.49) in the male-only corpus. There were not enough studies that reported hip fracture in the male-only corpus to pool results. The results of the analyses of fractures assessed at 12 and 24 months show similar findings as those assessed at the end of study. There was no evidence of statistical heterogeneity for any analysis.

Table 6. Summary of meta-regressions

	Coefficient	95% CI
MALE/FEMALE CORPUS, MALE PROPORTION EFFECT		
Any fracture	-0.18	-1.07, 0.70
Morphometric vertebral fracture	-0.35	-1.62, 0.93
MALE-ONLY CORPUS POPULATION		
Age effect		
Any fracture	0.02	-0.08, 0.13
Morphometric vertebral fracture	-0.10	-0.31, 0.12
Body mass index effect		
Any fracture	-0.001	-0.83, 0.83
Morphometric vertebral fracture	-0.38	-1.88, 1.12
T-score at lumbar spine effect		
Any fracture	-0.10	-1.53, 1.33
Morphometric vertebral fracture	-0.57	-7.24, 6.10
T-score at femoral neck effect		
Any fracture	-0.08	-3.66, 3.49
Morphometric vertebral fracture	8.42	-6.18, 23.03
T-score at total hip effect		
Any fracture	0.10	-17.93, 18.12
Morphometric vertebral fracture	2.91	-34.07, 39.88

Key: CI – Confidence interval

Effects of clinical risk factors on efficacy of bisphosphonates

In meta-regressions of male sex (in the male/female corpus) and baseline age, BMD, and BMI (in the male-only corpus), there was no evidence that differences in baseline clinical risk factors influenced the efficacy of bisphosphonates. There was a trend toward greater efficacy in males with higher BMD T-scores and lower BMI, but the associations were not statistically significant. The magnitude of effect of bisphosphonates did not vary with the change in proportion of males across studies, and we did not find any significant relationship between the other factors and effect size of bisphosphonates on male fracture. A summary of the results of meta-regressions is given in Table 6.

Summary

Results of our systematic review and meta-analysis showed that, for fracture endpoints that had been sufficiently studied, bisphosphonates have good efficacy in men. Our findings also failed to show and evidence that the treatment effect varies with baseline clinical risk level.

COMPARING FRACTURE ABSOLUTE RISK ASSESSMENT TOOLS: AN OSTEOPOROSIS CLINICAL INFORMATICS TOOL TO IMPROVE IDENTIFICATION AND CARE OF MALES AT HIGH FRACTURE RISK

To help clinicians identify male veterans at high fracture risk who had not received sufficient therapy for osteoporosis, we incorporated our VA-FARA tool into an electronic clinical informatics osteoporosis dashboard tool in VISN 21 during 2012. We compared the performance of the VA-FARA to an electronic version of the World Health Organization's (WHO) FRAX (eFRAX) using BMI rather than BMD. [40] A summary of the variables from the VA-FARA and the eFRAX, as operationalized in the VISN 21 data warehouse datasets, is given in electronic Appendix A (available from the principal investigator [PI] upon request).

Methods

Data source and dashboard development

An osteoporosis dashboard tool was developed using EMR data from the VISN 21 data warehouse, which stores near real-time transactional data recorded in the patient's EMR, including outpatient prescription information and clinic encounters, problem list diagnoses, laboratory values, vital signs, data collected in clinical reminders, and demographics.

Study design and patients

Using retrospective data, we compared the sensitivity of the VA-FARA versus the eFRAX for predicting fractures among cases with a first fracture. The cases were then matched to controls to compare the odds of fracture among patients classified as high risk by either tool. Cases were all male veterans aged ≥50 years with a first fragility fracture during the follow-up period from January 1, 2002 through June 25, 2013. January 1, 2002 was chosen for the study period start to allow up to 3 years observation prior to the fracture event to capture risk factor information. Fragility fracture was defined as an ICD-9 code for a fracture of the hip, spine, forearm, or proximal humerus. Male veterans without history of fracture who had encounters in VISN 21 were identified as controls and matched 1:1 to cases on age (within 1 year) and encounter date (within 1 month). An index date was defined as the date of the fracture event in the cases and the same date in the matched controls. For both groups, risk scores were calculated based on the risk factors present 3 years prior to and including the index date, excluding the outcome fracture from the risk score calculation.

eFRAX and VA-FARA scores

Patient risk scores for absolute fracture risk were calculated based on information obtained from the EHR. Risk factors were used to calculate eFRAX scores and are in electronic Appendix A (available from the PI upon request).[41] After identifying the presence of risk factors for patients, FRAX-based 10-year absolute fracture risks were obtained using risk-estimation tables available at http://www.shef/ac.uk/FRAX/charts.aspx. In cases where data was not available for a risk factor, such as the absence of information on family history of fracture, we assumed the factor was not present, but we conducted sensitivity analyses in which we assumed everyone with missing information on family history of fracture had that additional clinical risk factor. VA-FARA 10-year probabilities were calculated using the methods we described previously,[24] also summarized in an electronic Appendix A (available from the PI upon request).

Statistical analysis

Patients were classified as high-risk or non-high-risk by the two tools at the 20% cut-point for any major fracture, the 3% cut-point for hip fracture, and the combination of either hip or any major fracture at these cut-points, which is consistent with NOF guidelines for treatment in the US.[5] For the comparison of the tools' sensitivity to identify high-risk patients, cases were analyzed alone. Descriptive statistics were used for patient characteristics among cases and controls. *P*-values were calculated using Pearson's Chi-square test. Differences in sensitivity (percentage of fracture

patients classified as high-risk) were compared using a Chi-square test at 80% power and an alpha level of 0.05. Because of the study design, we were unable to directly compare specificity, positive predictive value, or negative predictive value for VA-FARA versus eFRAX. However, an odds ratio (OR) for high-risk classification with each tool was calculated along with the 95% confidence intervals (CI).

Results

Patients

A total of 544,040 males ages 50+ years received care in VISN 21 during the study period.[42] Of these, 3,475 (0.6%) had a documented traumatic fracture and were excluded. Of the remaining 540,565 patients, 7,145 (1.3%) had a documented fragility fracture, of which 4,730 (61.2%) were age-matched to at least 1 control. Of the 533,420 controls with no fragility fracture, 78,515 were eligible to be matched to a case. One control for each case was randomly selected, resulting in a final sample of 8,740 cases and controls. The mean (SD) age of the cohort was 67.0 (11.1) years; 65.8% were Caucasian and 27.8% were Black. As would be expected, most characteristics differed significantly between fracture and non-fracture patients, with cases having significantly more fracture risk factors compared to controls including white race, lower BMI, smoking and alcohol use, malnutrition, diabetes complications, fall risk, anticonvulsant use, proton pump inhibitor use, opioids, chronic steroid exposure, rheumatoid arthritis or depression diagnosis, and history of stroke.

Sensitivity: Performance of the tools for classifying cases as high risk

A summary of the numbers of cases classified as high risk using the two tools is provided in Table 7. Using the combined rule (i.e., exceeding the hip fracture rule at a threshold of 3% or the any major fracture rule at a threshold of 20%), the VA-FARA correctly classified 1,754 case patients as high risk compared to 761 with the eFRAX (40.1% versus 17.4%; p<0.001 using Pearson's Chi-square). By combining the two tools, the sensitivity could be increased to 44.8% (1,959 patients correctly classified as high risk using either tool), but 55.2% of fracture patients (2,411) would not be identified by either tool.

Table 7. Accuracy of VA-FARA versus eFRAX-based algorithm for correctly classifying fracture patients as high risk with hip fracture rules at a cutpoint of 3% and/or any major fracture rules at a cutpoint of 20%

	Sensitivity, fracture patients correctly classified				
	By eFRAX		By VA-FARA		
	N	%	N	%	*p*[a]
Hip fracture rule	761	17.4	1441	33.0	<0.001
Any major fracture rule	9	0.2	1510	34.6	NS
Hip or any major fracture	761	17.4	1754	40.1	<0.001

[a] *For eFRAX-based algorithm versus VA-FARA using Pearson's Chi-square for cell values >5 and Fischer's exact Chi-square for cell values ≤5*

Odds of fracture associated with high-risk classification

The odds of fracture associated with a high-risk classification with the two tools are given in Table 8. A patient who exceeded the 3% or 20% absolute risk thresholds for hip or any major fracture, respectively, with the VA-FARA tool was at a 35% increased risk of fracture (OR 1.35 95% CI 1.23, 1.47). For the eFRAX tool, the risk was increased by 17% (OR 1.17 95% CI 1.95, 1.32).

Table 8. Risk classifications and odds of fracture in patients classified as high risk versus not high risk using VA-FARA or eFRAX with hip fracture rules at a cutpoint of 3% and/or any major fracture rules at a cutpoint of 20%

	Cases						
	Correctly classified as high risk		Incorrectly classified as not high risk		OR[a]	*p*	95% CI
	N	%	N	%			
VA-FARA algorithm							
Hip fracture rule	1441	33	2929	67	1.21	<0.01	1.11, 1.33
Any major fracture rule	1510	34.6	2860	65.4	1.6	<0.01	1.45, 1.75
Hip or any major fracture[b]	1754	40.1	2616	59.9	1.35	<0.01	1.23, 1.47
eFRAX-based algorithm							
Hip fracture rule	761	17.4	3609	82.6	1.21	<0.01	1.08, 1.36
Any major fracture rule	9	0.2	4361	99.8	3	NS	0.75, 17.26
Hip or any major fracture[b]	761	17.4	3609	82.6	1.17	<0.01	1.05, 1.32

[a] *Odds ratio of fracture in patients classified as high risk using the specified rule.*
[b] *For joint rule, cases were classified as high risk if they exceeded either the 3% threshold on the hip fracture rule or the 20% threshold on the any major fracture rule. Controls were classified as non-high risk if they exceeded neither threshold for the two rules.*

Summary

We were successful in developing a computerized clinical decision support tool that provides clinicians timely information about patients' absolute fracture risk, which is important in making treatment decisions. In ongoing work we are measuring clinician response to the tool and conducting usability testing (see below under OTHER/ONGOING ANALYSES). However, initial findings from focus groups and informal discussions with users of the tool, we observed that there was considerable

uncertainty amongst clinicians about how best to use the information. That is, they were unsure whether they should treat on the basis of absolute risk or whether they should use the information provided in the dashboard to justify ordering a dual-energy x-ray absorptiometry (DXA) BMD scan. However, depending on which threshold of absolute risk was used on the VA-FARA, up to 40% of patients might require a DXA, which could have a substantial budget impact.

CONCLUSIONS

We succeeded in developing an absolute risk assessment tool that adequately discriminates between patients at high risk for fracture and those not at high risk (the VA-FARA). We determined that the tool is more effective than an analogous tool based on the FRAX (eFRAX). In cost-effectiveness analysis, we found that using VA-FARA to identify high-risk patients and then treating them on the basis of that risk, rather than on the basis of BMD T-score results, was more effective and less costly than other clinical decision strategies if bisphosphonates have any efficacy in non-osteoporotic, high-risk men. We also showed in meta-analysis that there is no evidence to support an inference that bisphosphonates have reduced efficacy in non-osteoporotic high-risk men.

OTHER/ONGOING WORK

Cognitive and motivational barriers to identifying and treating osteoporosis among US clinicians: We conducted 6 focus groups with 27 primary care and specialty clinicians at Salt Lake City VA, Richmond VA, and University of Utah Hospitals and Clinics. The goal was to identify information constructs that supported clinician decision-making surrounding osteoporosis, including cues for an index of suspicion that osteoporosis or fracture risk was present as well as information constructs that led a clinician to initiate treatment once an index of suspicion was present. Analyses are ongoing; preliminary results show numerous factors. The most interesting finding so far was the high level of uncertainty amongst clinicians about what to do in the presence of osteoporosis or fracture risk, how to weight different elements of the clinical picture, and whether to initiate bisphosphonates in the absence of certainty that the risks of treatment outweigh the benefits in male patients.

Salt Lake City's Bone Health Team: We have implemented a Bone Health Team (BHT) in Salt Lake City to provide osteoporosis screening and treatment in high-risk urban and rural patients in the Salt Lake City VA catchment area. The BHT has evaluated more than 975 individuals. In a comparison of patients who were screened by the BHT versus those who were not, we found that the BHT increased the likelihood that a patient would receive a BMD scan by 77 times, the likelihood that osteoporosis would be diagnosed by 27 times, and the likelihood that a bisphosphonate would be prescribed by 9 times.

Usability of an osteoporosis clinical support dashboard: In ongoing work, we are conducting usability testing with pharmacists and case managers who will interact with patient records to identify areas for further refinements to the tool using qualitative techniques.

Evaluation of clinician response to an osteoporosis clinical support dashboard: In ongoing work, we are analyzing the impact of the osteoporosis clinical dashboard tool before and after its implementation in 2012. Outcomes include incidence rates of BMD scan orders and bisphosphonate prescriptions.

Comparison of fracture risks and risk factors in Veteran compared to non-Veteran women: To determine if osteoporosis should be a greater health concern in women Veterans compared to non-Veterans, we compared fracture rates and BMD for Veterans and non-Veterans using Women's Health Initiative data. In this cohort study, participants were women ages 50-79. Outcomes were hip, central body, and limb fractures and hip, spine, and whole body BMD. Cox Proportional Hazards models were used to examine fracture rates for Veterans compared to non-Veterans. Baseline FRAX scores showed Veterans had higher 10-year probabilities for any major fracture (13.3 vs. 10.2; $p<0.01$) and hip fracture (4.1 vs. 2.2; $p<0.01$) compared to non-Veterans. The age-adjusted rate of hip fracture per 1000 person-years for Veterans was 3.3 versus 2.4 for non-Veterans ($p<0.01$). After adjustment, the hazards ratio for hip fracture was 1.24 (95% confidence interval 1.03-1.49) for Veterans versus non-Veterans. Hazards ratios at other anatomic sites did not differ by Veteran status. Mean BMD also did not differ by Veteran status at any site.

Patient-level meta-analysis of bisphosphonate efficacy in non-osteoporotic, high-risk males: To confirm our findings described above, we plan to conduct a patient-level meta-analysis to evaluate the efficacy of bisphosphonates in non-osteoporotic high-risk patients. Although our trial-level meta-analysis showed no reason to believe that there is reduced bisphosphonate efficacy in lower-risk patients, a patient-level analysis is needed to confirm these results.

The availability of clinical risk factor information in electronic medical records (EMRs): Many organizations are using their EMR data to identify patients at high risk for fracture using the FRAX calculator. We analyzed data in the University of Utah Hospitals and Clinics to determine whether this type of data was suitable for this use. Using data from the EMR, we found that only 10% of the cohort exceeded the NOF's 20% major fracture risk threshold and 32.5% exceeded the NOF's 3% hip fracture risk. Lower levels were observed when using BMI instead of BMD in the FRAX calculator. These findings suggest that using EMR data most likely underestimates the mean 10-year probability of any major fracture compared to other cohorts in the published literature. The difference may be in the nature of EMRs for supporting only passive data collection of risk factor information. This suggests that a FARA tool that is designed for use with EMR data is needed.

Using VA-FARA to identify men at high risk of fractures that are treated outside the VA: In ongoing analyses, we are evaluating the ability of VA-FARA to predict fractures that are treated outside the VA using Medicare data from the Centers for Medicare and Medicaid. Among the national cohort of male veterans, 80% of fractures are treated outside the VA. The current version of VA-FARA has only poor discrimination for identifying those fractures (C-statistics between 0.5 and 0.6). We are nearing the end of analyses the refine our absolute risk prediction algorithm.

Factors associated with screening or treatment initiation among male United States Veterans at risk for fracture: Male osteoporosis continues to be under-recognized and undertreated in men. We measured the association between a provider's osteoporosis identification and patient information constructs available at the time of each encounter. Using clinical and administrative data from the VHA system, we used a stepwise procedure to construct prognostic models for a combined outcome of osteoporosis diagnosis, treatment, or a BMD test order using time-varying covariates and Cox regression. We ran separate models for patients with at least one primary care visit and patients with only secondary care visits in the pre-index period. Some of the strongest predictors of clinical osteoporosis identification were history of gonadotropin-releasing hormone (GnRH) agonist exposure, fragility fractures, and diagnosis of rheumatoid arthritis. Several known clinical risk factors for fracture were not correlated with osteoporosis risk including smoking and alcohol abuse. Results suggest that clinicians are relying on some, but not all, clinical risk factors when assessing osteoporosis risk.

Using VA-FARA to identity female veterans at high risk of fracture: In ongoing analyses, we are evaluating the ability of VA-FARA to predict fractures that occur in women. Although the proportion of elderly women in the VA is small (fewer than 5%), this proportion is increasing, and so extending the ability of our tool to function in this population is essential.

The clinical epidemiology of male osteoporosis: Osteoporosis is now recognized as an important public health problem in men and remains largely underdiagnosed and undertreated. The objective of this paper is to provide an overview of recent findings in male osteoporosis, including pathophysiology, epidemiology, and incidence and burden of fracture, and discuss current knowledge about the evaluation and treatment of osteoporosis in males. In particular, clinical practice guidelines, fracture risk assessment, and evidence of treatment effectiveness in men are addressed.

Analysis of osteoporosis patterns and outcomes among postmenopausal veterans: Adherence and persistence with bisphosphonates are frequently poor, and stopping, restarting, or switching bisphosphonates is common. We evaluated bisphosphonate change behaviors (switching, discontinuing, or reinitiating) over time, as well as fractures and costs, among a large, national cohort of postmenopausal veterans. Most bisphosphonate patients discontinue treatment at some point, which did not significantly increase the risk of fracture in this majority non-high risk population. Bisphosphonate change behaviors were associated with significantly lower osteoporosis costs, but significantly higher total costs.

Calculating the baseline fracture incidence in non-risk patients: We developed a method for deriving the incidence for persons without risk factors by adjustment of incidence in the population (incidence_pop). We calculated incidence_no_risk using the relative risk for events due to risk factors (RR_risk), incidence_pop, and the prevalence of the risk factor (pRF), which are typically available in the literature. Since the incidence for patient with risk factors (incidence_risk) can be expressed as incidence_risk = incidence_no_risk × RR_risk, we found that incidence_no_risk = incidence_pop/((RR_risk × pRF) + (1 - pRF)). We validated the equation by modeling the fracture incidence in high-risk patients in an osteoporosis transition-state model. After adjustment of incidence_pop using incidence_no_risk as the baseline incidence, the model accurately predicted hip fractures (2.27/1000 patient-years).

Therefore, incidence_no_risk can be calculated using this method based on the event incidence for the study population, the relative risk increase associated with the risk factor, and the prevalence of the risk factor.

Sources of care for Veterans who fracture: We examined the relationship between distance to the nearest Veterans Affairs Medical Center (VAMC) and care site for elderly male Veterans who were dually eligible to receive VA and Medicare services. We used a cross-sectional design to identify male Veterans (age ≥70) with a non-traumatic hip, wrist, or spine fracture from all male Veterans who received care in the VA between 2002 and 2008 with a recent history of VA utilization. Logistic regression was used to estimate the effect of distance on VA treatment for fracture while controlling for age, race, comorbidity index, region, and geographic characteristics. Patients that lived further from a VAMC were less likely to have their fracture treated at the VA than those who were located within 5 miles of a VAMC. Among elderly males with a history of VA system utilization, greater distance to the nearest VAMC was significantly associated with a decreased likelihood of using the VA for fracture treatment.

LIST OF PUBLICATIONS AND PRODUCTS

Published/in-press peer-reviewed journal articles

1. *LaFleur J*, McAdam-Marx C, Asche CV, Alder S, Sheng X, Nebeker J, Brixner DI, Silverman S. Clinical risk factors for fracture among postmenopausal patients at risk for fracture: A historical cohort study using electronic medical record data. Journal of Bone and Mineral Metabolism. 2011;29(2):193.
2. *LaFleur J*, Nelson RE, Yao Y, Adler R, Nebeker J. Validated risk rule using computerized data to identify males at high risk for fracture. Osteoporosis International. 2012:23(3):1017-27.
3. Nelson R, Nebeker J, Sauer B, *LaFleur J*. Factors associated with screening or treatment initiation among male United States veterans at risk for osteoporosis fracture. Bone. 2012; 50(4):983-8.
4. Nelson SD, Nelson RE, Cannon GW, Lawrence P, Battistone MJ, Grotzke M, Rosenblum Y, *LaFleur J*. Cost-effectiveness of training rural providers to identify and treat patients at risk for fragility fractures. Osteoporosis International. 2014; 25(12):2701-7.
5. Willson T, Nelson S, Newbold J, Nelson RE, *LaFleur J*. The clinical epidemiology of male osteoporosis: A review of the recent literature. Clinical Epidemiology. 2015;87:65-76.
6. Unni S, Gunning K, Curtis J, Yao Y, *LaFleur J*. An evaluation of clinical risk factors for estimating fracture risk in postmenopausal osteoporosis using an electronic medical record database. Osteoporosis International. 2015; 26(2):581-7.
7. *LaFleur J*, Steenhoek CL, Horne J, Meier J, Nebeker JR, Mambourg S, Swislocki A, Carmichael J. Comparing fracture absolute risk assessment tools: An osteoporosis clinical informatics tool to improve identification and care of males at high fracture risk. Annals of Pharmacotherapy. 2015; 49(5):506-14.
8. *LaFleur J*, DuVall S, Willson T, Gintern T, Patterson O, Cheng Y, Knippenberg K, Hayden C, Adler R, Curtis J, Agodoa I, Stolshek B, Nelson RE. Analysis of osteoporosis treatment patterns and outcomes among postmenopausal veterans. Bone. 2015;78:174-85.
9. Colon-Emeric C, Pieper CF, Grubber J, Van Scoyoc L, Schnell ML, Van Houtven CH, Pearson M, *LaFleur J*, Lyles KW, Adler R. Correlation of hip fracture with other fracture types: Toward a rational combined hip fracture endpoint. Bone. 2015;81:67-71.
10. Nelson S, Malone D, *LaFleur J*. Calculating the baseline incidence in patients without risk factors: A strategy for economic evaluation. Pharmacoeconomics. 2015; In press.
11. Abstracts
12. *LaFleur J*, Smith J, Nelson S, Nelson R, Adler RA, Nebeker JR, Malone DC. Cost-effectiveness analysis of different strategies for fragility fracture prevention in United States veterans. Value in Health;15(4):A8. Also presented as a podium presentation at the International Society for Pharmacoeconomics and Outcomes Research 17th International Meeting. Washington, DC. June 2012.
13. Nelson S, Malone D, *LaFleur J*. Calculating the baseline fracture incidence in non-risk patients: A strategy for cost-effectiveness modeling. Value in Health 2014;17:A47. Also presented as a poster at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 19th Annual International Meeting. Montreal, Quebec, Canada. May-June 2014.
14. Nelson S, Willson T, *LaFleur J*. Where do Veterans receive their care for non-traumatic bone fractures? Value in Health 2014;17:A146. Also presented as a poster at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 19th Annual International Meeting. Montreal, Quebec, Canada. May-June 2014.

15. Cheng Y, Jiao T, Willson T, Reese T, Stoddard GJ, *LaFleur J*. Bisphosphonates for fracture prevention in males: A systematic review and meta-analysis. Value in Health. 2015;18(3):A154. Also presented as a poster at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 20th Annual International Meeting. Philadelphia, Pennsylvania. May 2015.
16. *LaFleur J*, Rillamas-Sun E, Colon-Emeric CS, Knippenberg KA, Ensrud KE, Gray SL, LaCroix AZ. Fracture rates and bone density among postmenopausal Veteran and non-Veteran women from the Women's Health Initiative. The Gerontologist. In press.

Posters

1. Unni S, Yao Y, Gunning K, Curtis J, *LaFleur J*. "Analysis of clinical risk factors for post-menopausal osteoporosis based on WHO fracture risk assessment (FRAX) tool." American Society of Bone and Mineral Research 2011 Annual Meeting. San Diego, California. September 2011.
2. *LaFleur J*, Ginter T, Hayden C, DuVall SL, Adler R, Nebeker J. "Bone mineral density screening and osteoporosis rates in veterans." American Society of Bone and Mineral Research 2011 Annual Meeting. San Diego, California. September 2011.
3. Tomic KS, *LaFleur J*, Palmer L, Smith DM, Paoli C, Agodoa I, Yurgin N. "Risk of Fracture among treated and untreated men with osteoporosis." American Society of Bone and Mineral Research 2012 Annual Meeting. Minneapolis, Minnesota. October 2012.
4. Kazis L, Lee A, Li M, *LaFleur J*, Vlad SC, Carey K, Chew P, Chandler D, Yurgin N, and Adler R. "Predictors of Non-Adherence to Bisphosphonates for Male Veterans with Osteoporosis and/or Osteoporotic Fracture: Importance of Mental Health Conditions." American Society of Bone and Mineral Research 2012 Annual Meeting. Minneapolis, Minnesota. October 2012.
5. Tomic KS, *LaFleur J*, Palmer L, Smith DM, Paoli C, Agodoa I, Yurgin N. "Risk of Fracture among treated and untreated men with osteoporosis." American College of Rheumatology Annual Meeting. Washington, DC. November 2012.
6. Luk E, Meier J, *LaFleur J*, Swislocki A, Carmichael J. "Pharmacist impact on collaborative osteoporosis care for veterans at high fracture risk." American Society of Health System Pharmacists Midyear Meeting. Las Vegas, NV. December 2012.
7. Steenhoek C, *LaFleur J*, Palmer L, Meier J, Swislocki A, Carmichael J. "Fracture risk assessment tool comparison in male veterans." American Society of Health-System Pharmacists Midyear Meeting. Las Vegas, NV. December 2012.
8. *LaFleur J*, Willson T, Hayden C, Adler B, Nebeker J. Validation of a fracture risk rule using computerized data to identify veterans at high risk for fracture. Western Pharmacoeconomics Conference. Tucson, Arizona. March 2013.
9. *LaFleur J*, Ginter T, Curtis J, Adler R, Agodoa I, Stolshek B, and Nelson R, DuVall S. A novel method for obtaining bone mineral densities from a dataset of radiology reports and clinic notes: Natural language processing in a national cohort of postmenopausal veterans. American Society of Bone and Mineral Research 2013 Annual Meeting. Baltimore, Maryland. October 2013.
10. *LaFleur J*, DuVall S, Curtis JR, Adler RA, Willson T, Agodoa I, Stolshek B, and Nelson RE. Association between Bisphosphonate Switching Behavior and Cost Outcomes in Postmenopausal United States Veterans. American College of Rheumatology 2013 Annual Meeting. San Diego, California. October 2013.
11. *LaFleur J*, Steenhoek CL, Horne J, Meier J, Mambourg S, Nebeker JE, Swislocki A, Carmichael J. Comparing fracture absolute risk assessment tools: An osteoporosis clinical informatics tool to improve identification and care of males at high risk of first fracture. American Federation for Medical Research Western Regional Meeting, Carmel, California, January 2014.
12. *LaFleur J*, Steenhoek CL, Horne J, Meier J, Mambourg S, Nebeker JE, Swislocki A, Carmichael J. Comparing fracture absolute risk assessment tools: An osteoporosis clinical informatics tool to improve identification and care of males at high risk of first fracture. Joint meeting of the International Society of Endocrinology and The Endocrine Society: ICE/ENDO 2014. Chicago, Illinois. June 2014.
13. Colon-Emeric C, Pieper CF, Grubber J, Van Scoyoc L, Schnell M, Houtven CV, Pearson M, *LaFleur J*, Lyles KW, Adler R. Correlation of other fracture types with hip fracture: Toward a rational combined hip fracture endpoint. American Society for Bone and Mineral Research 2014 Annual Meeting. Houston, Texas. September 2014.
14. Biksacky M, Crook J, Cheng Y, Knippenberg K, *LaFleur J*. Reasons for switching and discontinuing bisphosphonate therapy: A historical chart review of postmenopausal veterans. The 49th American Society of Health-System Pharmacists Midyear Clinical Meeting and Exhibition. Anaheim, California. December 2014.

15. Miller K, Grotzke M, Lawrence P, Rosenblum Y, Nelson RE, *LaFleur J*, Cannon GW. "Implementation of a bone health team markedly improves osteoporosis screening, diagnosis, and treatment initiation rates compared to standard primary care practice." American Society of Rheumatology and Association of Rheumatology Health Professionals Annual Meeting. San Francisco, California. November 2015. [Accepted]

Submitted manuscripts

1. *LaFleur J*, Nelson S, Malone D, Nelson R, Adler B, Nebeker J. Cost-effectiveness of strategies for male osteoporosis screening and fracture prevention using fracture absolute risk assessment PLOS One. Under consideration.

Other papers still in development